Warning Concerning Copyright Restrictions

The Copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyright material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction not be "used for any purposes other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

2

Classical Conditioning

Overview

Twenty-five years ago I ate crab for the first time. Shortly thereafter, I developed a horrific stomach flu (I will spare you the details) which had been going around. To this day I get nauseous when I am offered crab. This is an example of classical conditioning. As discussed in the previous chapter, a conditioned stimulus (CS—this case, crab) had been paired with an unconditioned stimulus (US—in this case, flu) that produced an unconditioned response (UR—in this case, extreme nausea). The result was that the CS acquired the ability to evoke a conditioned response (CR—again nausea in this case). As we will discuss later in this chapter, food aversion can be a particularly powerful version of classical conditioning. It certainly is in my case—I am getting nauseous just writing this paragraph.

The Phenomena of Classical Conditioning: Eye Blink in Humans

Most contemporary research on classical conditioning has been conducted on animals, but in the past a great deal of research was conducted on humans. A popular paradigm for studying human conditioning was eyeblink conditioning. Imagine what it would be like to be a subject in a procedure such as eyeblink conditioning. In a typical eyeblink-conditioning experiment, the subject is fitted with a padded headband containing a nozzle pointed toward the eye, and a wire measuring muscle activity in the eyelid is taped to the eyelid. The US is a puff of air directed toward the outside of the cornea, and the CS is a light or tone. The US normally evokes a UR of an eye blink, and conditioning is concerned with how the CS comes to evoke a similar CR. Figure 2.1 shows a diagram describing this situation much as Figure 1.3 describes Pavlov's conditioning of salivation in dogs. However, in this case we are using the diagram to represent the situation of human conditioning of the eye blink.

In human eyeblink conditioning, one can demonstrate all of the basic phenomena of classical conditioning. For instance, consider acquisition and extinc-

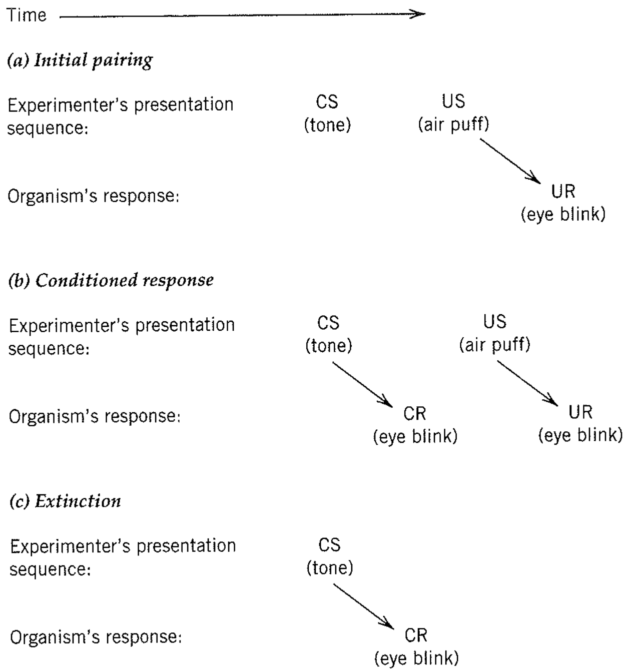

FIGURE 2.1 Experimental procedure for classical conditioning of the eye blink. Compare with Figure 1.3.

tion of the conditioned response. In one experiment (Moore & Gormezano, 1961), the interval between the CS and the US was 500 msec. Subjects were given 70 acquisition trials in which the CS was followed by the US and then 20 extinction trials in which the CS was presented alone. Figure 2.2 shows the percentage of tri-

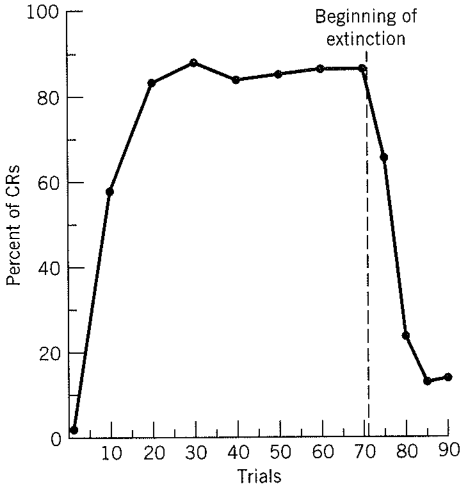

FIGURE 2.2 Probability of a conditioned eye blink during 10-trial blocks of acquisition and 5-trial blocks of extinction. (From Moore & Gormezano, 1961.)

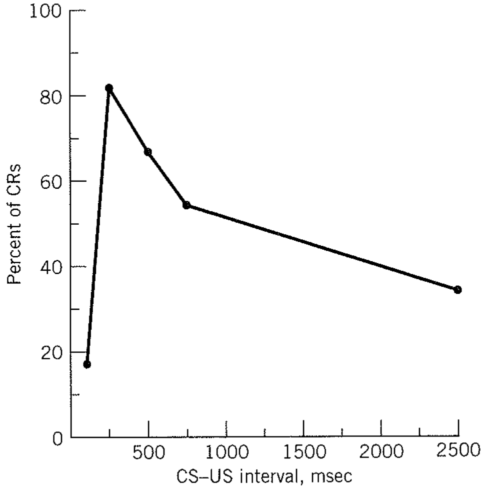

FIGURE 2.3 Percentage of conditioned eye blink as a function of CS–US interval. (From McAllister, 1953.)

als in which subjects emitted a CR in anticipation of the puff of air. As in the salivation data presented in Figure 1.4, these human subjects showed a rather standard conditioning curve, in which the probability of the CR increased, followed by a standard extinction curve, in which the probability of the CR decreased. If one waits for a while after extinction, one finds spontaneous recovery of the extinguished CR. That is, when tested again without the US, the probability of the CR increases from its former extinguished level (Grant, Hunter, & Patel, 1958).

Another standard parameter of a classical conditioning experiment is the interval between the CS and the US. McAllister (1953) varied the time between a CS of a tone and the US of an air puff using intervals of 100, 250, 450, 700, and 2500 msec. Figure 2.3 shows the results in terms of percentage of conditioned responses after 20 conditioning trials. Nearly maximal conditioning was achieved in the intervals between 250 and 700 msec, which are typical values for optimal conditioning in many, but not all, conditioning paradigms.

The classical conditioning of the human eye blink demonstrates typical acquisition and extinction functions as well as typical effects of CS–US intervals.

Sensitization and Habituation

Eyeblink conditioning can be used to show two other learning phenomena that should be distinguished from true classical conditioning. The first is **sensitization**. In the design of Figure 2.1 the tone always precedes the puff of air. Figure 2.3 showed that subjects were sensitive to the exact interval between tone and air puff. What happens, however, if there is no relationship between the CS and US? In such an experiment, a subject will be exposed to a series of air puffs and hear a series of tones, but there will not be any relationship between them. Nonetheless, the subject will show some increase in the tendency to show a CR of blinking when the tone occurs (Grant & Norris, 1947). Indeed, the experiment can be run with the US presented alone a number of times before the CS is first given. Then, upon first occurrence of the CS there is an increased tendency to emit an eye blink. Thus, the mere presence of the US increases the subjects' tendency to respond with eye blink to any stimulus. This occurs in many classical conditioning paradigms and indicates that the US has in some sense made the subject more sensitive to any stimulus. It is as if the subject has become "jumpy." This differs from classical conditioning because it does not depend on any relationship between CS and US in conditioning. Well-designed studies of classical conditioning often include a control group where there is no relationship as well as an experimental group where the CS does precede the US. If the experimental group shows a larger CR, then we can assume that some actual associative learning, rather than just a sensitization process, has taken place.

The other nonassociative learning phenomenon is called **habituation**. If the US is repeated over and over again, the magnitude of the UR that it evokes often tends to decrease. Such habituation has been reported in eyelid conditioning experiments (e.g., Beecroft, 1966; Oldfield, 1937). In this case, the magnitude of the response is decreasing as the subject becomes exposed to the US. It is as if the subject is becoming "toughened" and no longer has as strong a reaction to the US. As a consequence, the size of the conditioned response can decrease. For instance, Grant (1939) found a gradual decrease of both the frequency and amplitude of the eyelid response to the CS.

Both sensitization and habituation are classified as *nonassociative* forms of learning because they do not depend on the relationship between the CS and the US. Note that they pull in opposite directions. Sensitization results in increased response, while habituation results in decreased response. From experiment to experiment, their net effect may be either an increased responsiveness or a decreased responsiveness (Groves & Thompson, 1970). In all cases, it is necessary to separate these nonassociative forms of learning from the associative learning that defines classical conditioning.

Two other forms of learning are sensitization, which is increased responsiveness to many stimuli because of US exposure, and habituation, which is decreased responsiveness to the US because of US exposure.

Conditioning and Awareness

The conditioned eye blink is adaptive. By blinking in anticipation of the puff of air, the subjects are protecting themselves from an aversive stimulus. The fact that the eye blink is adaptive has raised the issue of whether human subjects are voluntarily choosing to blink so that they can avoid the light. Human subjects report being aware of the CS–US relationship and blinking in response to the US (Grant, 1973). However, they tend to be unaware that they are blinking to the CS in anticipation of the US. It has nonetheless been argued that certain eyeblink CRs look like eye blinks that subjects give when instructed to blink; in particular, the eye closes more sharply, more rapidly, and stays closed longer when it is a voluntary response (Spence & Ross, 1959). Considerable controversy remains over whether it is possible to discriminate between automatic and voluntary eye blinks (Gormezano, 1965; Ross, 1965).

At one time, classical conditioning was considered automatic, and instrumental conditioning was considered voluntary. This viewpoint led to the argument that the purported voluntary eye blinks in a classical conditioning paradigm were really instances of covert instrumental conditioning in which the human was responding for the reward of avoiding the aversive puff of air. This distinction has not proved useful and is not pursued here. Whether voluntary or automatic, certain behavioral regularities tend to be associated with conditioning. Research has focused on understanding these regularities. An important theme in modern research is that classical conditioning serves an adaptive function, whatever the awareness of the organism or the degree of voluntariness in responding.

Classical conditioning shows similar behavioral properties whether or not the subject is aware.

What This Chapter Covers

One of the striking features of classical conditioning is its ubiquity. Virtually all organisms can be conditioned. Dogs and humans have already been mentioned. This chapter reviews classical conditioning in organisms ranging from sea slugs to rabbits, as well as some of the enormous variety of stimuli that can be used for the CS and the US.

Classical conditioning is often considered the paradigm of choice for studying how associations are made. The similar properties it displays over a wide range of situations and organisms might lead to the conjecture that learning is taking place according to the same neural mechanisms in all these situations. Recent research on the neural basis of some instances of classical conditioning in certain organisms is discussed next. The research shows that different neural mechanisms underlie classical conditioning in different organisms and, indeed, that different mechanisms underlie different types of conditioning in the same organism. Creatures have found various ways of forming associations in different situations. Constant across these circumstances is the need to form associations. The constancy in the behavioral manifestations of classical conditioning reflects the constancy of this need.

After this review, the chapter discusses the behavioral properties of classical conditioning. In particular, subsequent sections address the following questions:

What is associated?
What is the conditioned stimulus?
What is the conditioned response?
What is the nature of the association?

The answers are surprising in light of earlier views about classical conditioning. The Rescorla–Wagner theory, an elegant but simple theory that captures much of the complex structure of the data, is presented later in this chapter.

At a behavioral level, a wide variety of organisms display similar classical conditioning phenomena when a neutral CS is followed by a US.

Neural Basis of Classical Conditioning

As reviewed in Chapter 1, Pavlov speculated, incorrectly, about what was happening in the nervous system to underlie classical conditioning. More recently, researchers have traced the neural bases of certain instances of classical conditioning in certain organisms. This research offers a glimpse of the neural mechanisms behind the behavioral phenomena and may provide a sobering influence on overly simplistic or overly grandiose interpretations of classical conditioning.

As we have seen, neural information processing takes place through transmission of signals among neurons, which are the individual cells that make up the nervous system. Neurons transmit signals from one part of the nervous system to another by sending electrical pulses along their axons. The axon of one neuron makes contact with the cell bodies of other neurons. The point of contact between the axon of one neuron and the cell body of another neuron is called the *synapse*, and communication is achieved by transmitter chemicals, called *neurotransmitters*, going from the axon of one neuron to the other neuron. These neurotransmitters can increase or decrease the electrical potential of the neuron. If enough electrical potential accumulates on its cell body, a neuron sends a signal down its axon. It is generally believed that learning involves changes in the effectiveness of synaptic connections among neurons, so that one neuron comes to produce greater changes in the electrical potential of another.

All learning, including classical conditioning, involves changes in the effectiveness of synaptic connections among neurons.

Simple Learning in *Aplysia* (Sea Slug)

Some of the most influential research on neural mechanisms has been done on invertebrates that have simple nervous systems with very large neurons. The simplicity of their nervous systems allows researchers to understand in detail how the systems work; the large neurons make studying what is happening to an individual neuron relatively easy.

One of the creatures that has received extensive study is the sea slug, ***Aplysia*** *californica* (see Figure 2.4). The gill (the respiratory organ of the *Aplysia*) and the siphon (a fleshy spout, surrounding and enveloping the gill, used to expel water) of *Aplysia* have withdrawal reflexes that can be evoked by touching the siphon (or other nearby parts, such as the mantle). This reflex is largely controlled by direct synaptic connections between the sensory neurons that are excited by the tactile stimulation and the motor neurons that control the reflex. Figure 2.5 shows schematically the synaptic connections for the gill-withdrawal reflex. The sensory neuron from the siphon skin synapses directly onto the motor neuron for withdrawing the gill. Thus, touching the siphon stimulates the sensory neuron, which stimulates the motor neuron, which in turn evokes the gill-withdrawal reflex.

The withdrawal reflex to tactile stimulation is not very strong and tends to weaken with repeated touching. The strength of the response can be enhanced by a classical conditioning procedure that pairs tactile stimulation to the siphon (CS) with shock to the tail (US). After five such pairings, the tactile stimulation (CS) evokes a much stronger withdrawal reflex than is present without pairing with US. Carew, Hawkins, and Kandel (1983) compared using a CS of tactile stimulation to the siphon with a CS of tactile stimulation to the mantle. If the siphon stimulus is paired with the US, it evokes a stronger withdrawal reflex than that evoked by the mantle stimulus; if the mantle stimulus is paired with the US, it comes to evoke a stronger reflex than that evoked by the siphon stimulus. Thus, *Aplysia* can learn to discriminate between the two locations.

The mechanism of conditioning appears to involve the facilitating interneurons shown in Figure 2.5. These are neurons onto which sensory neurons form the tail synapse. They, in turn, synapse on the axons of the sensory neurons from the siphon, providing an example of an **axoaxonic synapse**—a synapse on a synapse. The shock to the tail activates these interneurons, which operate on the synaptic connection between the sensory neurons from the

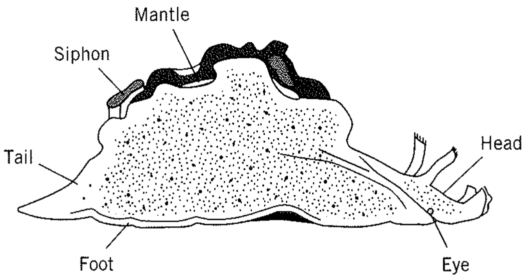

FIGURE 2.4 The *Aplysia*. *Source:* From *Cellular basis of behavior* by Eric Kandel. Copyright © 1976 by W. H. Freeman and Company. Reprinted with permission.

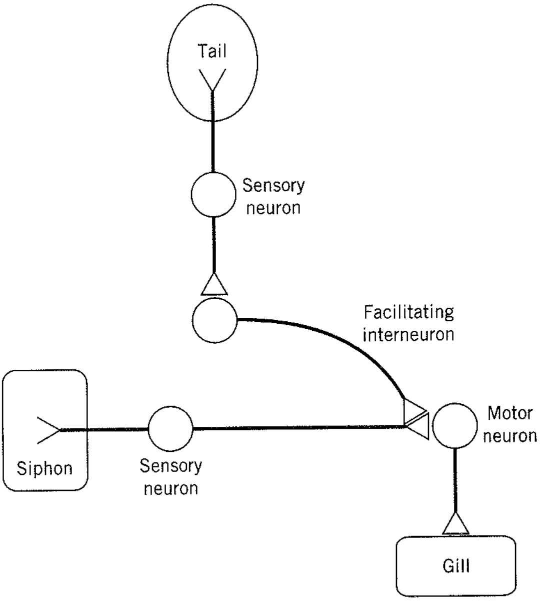

FIGURE 2.5 Neural connections underlying the conditioning of the gill-withdrawal reflex in *Aplysia.*

siphon and the motor neurons. They change the synaptic connection by increasing the release of the neurotransmitter from the sensory neuron coming from the siphon. This process is referred to as **presynaptic facilitation**, because it enhances a process that is occurring on the axon side of the synapse.

Bailey and Chen (1983) studied the structural basis for these changes in neurotransmitter release. When there was conditioning, they found an increase in sites on the axon where neurotransmitter release occurs. A neurotransmitter called *serotonin* released by the facilitating interneurons starts a relatively long chemical chain reaction that results in more release sites at the synapse of the sensory neuron from the siphon. This presynaptic facilitation is maximal if the sensory neuron from the siphon is active just before the interneuron from the tail. If the sensory neuron from the siphon (CS) has just fired, the chemical chain reaction is enhanced, resulting in the exposure of release sites at the synapse between the sensory neuron and the motor neuron (Hawkins, Abrams, Carew, & Kandel, 1983). This dependency of presynaptic facilitation on recent activity from the siphon makes classical conditioning in *Aplysia* dependent on CS–US intervals. Maximal conditioning occurs if the tactile stimulation (CS) takes place 0.5 sec before the shock (US) and no conditioning occurs after as little as 2 sec. The axon from the siphon is only briefly in the state that results in presynaptic facilitation.

Classical conditioning has been studied in another invertebrate, the nudibranch mollusk, *Hermissenda crassiconis* (Alkon, 1984). Classical conditioning in this animal also involves a change in neurotransmitter release, but the actual cellular mechanisms differ from those at work in *Aplysia*. These studies show that there is no single mechanism of learning at the cellular level.

Classical conditioning in Aplysia *is produced by enhancing the neurotransmitter release between the sensory neuron from the CS and the motor neuron producing the CR.*

Classical Conditioning of the Eye Blink in the Rabbit

Thompson (e.g., 1986) engaged in an extensive project studying the neural basis for a more complex example of classical conditioning, the conditioning of the eye blink in the rabbit. In the standard procedure, the US is a puff of air to the eye, which produces a UR of an eye blink (Gormezano, Kehoe, & Marshall, 1983). The CS is a tone, which also comes to evoke an eye blink (CR), although at greater latency. (The latency for the UR is about 20 msec from air puff to blink, whereas the latency for the CR is about 70 msec from tone to blink.)

Figure 2.6 illustrates some of the complex circuitry that is relevant to understanding this instance of classical conditioning. Sensory neurons from the cornea synapse onto the fifth cranial nerve, from which neurons go to the sixth and seventh cranial nerves, from which motor neurons go to produce the eye blink. This

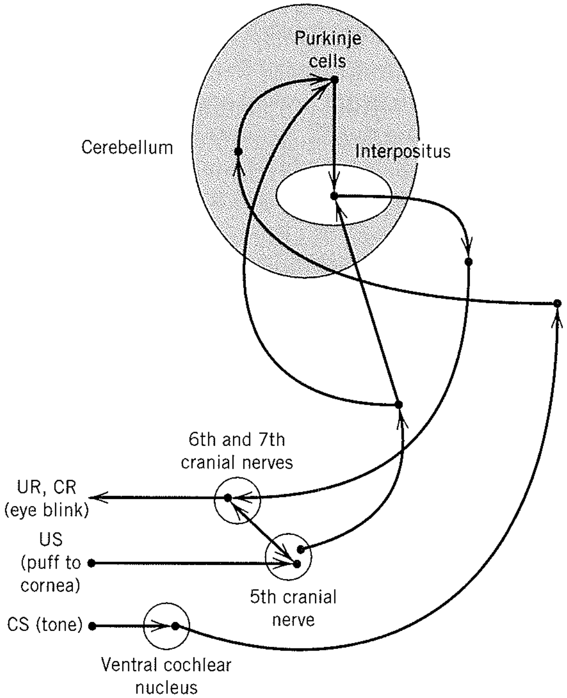

FIGURE 2.6 Simplified diagram of the neural circuitry responsible for eye blink conditioning.

circuit, which produces the UR, takes about 20 msec. A second, longer circuit goes from the fifth cranial nerve to an intermediate neuron, and from there to the cerebellum (a subcortical structure—see Figure 1.15). Among other neurons in the cerebellum, this path synapses onto cells in a structure called the interpositus. There is also a path of synapses going back from the interpositus to the eye blink, which produces a discrete electromyographic[1] component to the eye blink (UR) about 70 msec after the puff of air (US). The longer latency reflects the greater number of neurons along this path (at least six).

The longer US–UR circuit is of interest because it is involved in the conditioning. There is a circuit by which sensory neurons encoding the tone (CS) synapse onto the ventral cochlear nucleus, and from there by intermediate neurons to Purkinje cells in the cerebellum. Conditioning can be produced by directly stimulating this CS circuit and omitting the tone. It is generally thought that the conditioning is taking place in the cerebellum, since lesions (selective removal of neural tissue) to the cerebellum eliminate eyeblink conditioning. In addition to the paths by which the CS projects to the Purkinje cells, there are paths by which the US projects to Purkinje cells. The Purkinje cells synapse onto the interpositus, which is part of both the CR and the UR paths. The logic by which the CS is thought to evoke the CR is somewhat complicated:

1. The Purkinje cells normally tend to inhibit the interpositus from evoking the response.
2. Learning involves developing inhibitory connections from the CS path to the Purkinje cells.

It is proposed that learning involves inhibiting the Purkinje cells, which normally inhibit the eye blink, thereby enabling the CR. (In effect, two neural negatives are combining to make a neural positive.) Electrodes put into the cerebellum to record from the Purkinje cells show reduced firing after conditioning, which is in line with this proposal. This situation is different and more complex than that of conditioning in *Aplysia*. The path of conditioning in *Aplysia* involves two neurons directly associated so that the sensory neuron turns on the motor neuron. In the rabbit, the path involves more than a half-dozen neurons and has the sensory stimulus turning off the cells that normally turn off the eye blink. Although not as thoroughly studied, other instances of mammalian classical conditioning appear equally or more complex and involve different neural structures from the spinal cord to the cortex and different neural circuitry and modifications (Thompson, Donegon, & Lavond, 1988). No single neural process underlies all classical conditioning.

Despite the diversity of neural realizations, classical conditioning presents a rather consistent picture at the behavioral level. These behavioral regularities are identified next.

---

[1] Electromyographic recordings measure electric activity associated with contraction of muscles.

Classical conditioning of the eye blink takes place in the cerebellum, where paths involving the CS, the US, and the CR meet, and it appears to involve learning to inhibit inhibiting connections.

S–S or S–R Associations?

Every time I pick up the leash, my dog bounds to the door. This is an instance of classical conditioning in which the CS is the leash, the US is starting a walk, and the CR and UR are going to the door. An interesting question that has occupied researchers is, what is the CS (leash) associated to? Figure 2.7 shows two possible answers to this question. Figure 2.7*a* illustrates the more obvious answer, given the experimental description: from repeated pairings of CS and UR, the CS comes to evoke the UR. The CR then is just a version of the UR. Thus, my dog is going to the door because that is what he does when a walk begins. This association is called a stimulus–response or **S–R association**. Another possibility is that the CS becomes associated to the US (Figure 2.7*b*). The organism behaves as if it expects the CS to be followed by the US; it emits the CR in anticipation of the US and perhaps as preparation for the US. Thus, my dog is going to the door because the leash is signaling that a walk will begin. This situation is called a stimulus–stimulus, or **S–S association**.

In reviewing conditioning in *Aplysia*, the choice between the S–S position and the S–R position may seem obvious. The sensory neuron encoding the CS from the siphon was directly connected to the motor neuron that produced the gill retraction, a clear victory for the S–R position. However, *Aplysia* has a nervous system very different from that of mammals. In the case of the circuitry underlying the eye blink in the rabbit, the issue is far from clear. The CS, US, and

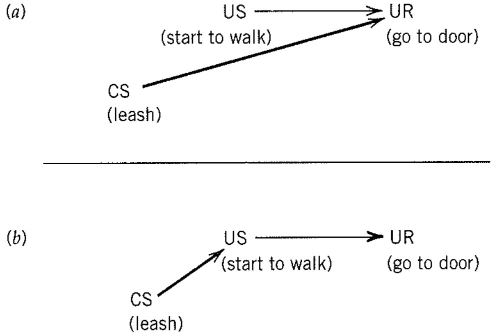

FIGURE 2.7 The two possibilities for association formation in classical conditioning: (*a*) the CS is associated to the UR, and (*b*) the CS is associated to the US.

UR all met in the cerebellum, it could be argued that the CS and the US were associated at the Purkinje cells.

A peculiarity of the S–S position from a biological point of view is evident in the sentence,"The organism comes to behave *as if it expected* the CS would be followed by the US."Although human subjects may have conscious expectations, ascribing such conscious thoughts to animals, such as rabbits, is problematic. Hence, the words *as if* are used. Some rather simple associative mechanisms can produce behavior that mimics conscious deliberation. One example is provided by the Rescorla–Wagner theory, which is reviewed later in this chapter.

Although the neural evidence is unclear, several behavioral paradigms provide information relevant to deciding between an S–S position and an S–R position. These include response prevention, US devaluation, sensory preconditioning, and second-order conditioning. Each of these paradigms is considered in the following subsections.

The two possibilities for the association are the S–R hypothesis that the CS becomes associated to the UR and the S–S hypothesis that the CS becomes associated to the US.

Response–Prevention Paradigm

One way to test the S–R position is to block the response and determine whether associations can still be formed. This method is referred to as the **response–prevention paradigm**. For instance, Light and Gantt (1936) paired a CS of a buzzer with a US of shock to dogs' paws. This US normally produces a UR of leg lifting, which gets conditioned to the CS. However, they prevented this response by temporally damaging the spinal motor nerves that went to the affected legs. Thus, the dogs never had an opportunity to perform the response and so to form an association between CS and UR. The spinal nerves recovered, and within a couple of months the dogs had complete use of their limbs. Then when presented with the CS of a tone, they displayed the CR of leg lifting. More recent studies have shown similar conditioning when an animal was prevented from responding by use of drugs (Fitzgerald, Martin, & O'Brien, 1973) or by temporary paralysis with curare (Leaf, 1964). In all these experiments, there is a US but no UR to which the CS can become associated. The existence of conditioning supports the S–S position over the S–R position.

It could be argued that this interpretation of the S–R association is too peripheral. Even if these organisms were prevented from executing motor responses, their central nervous systems might well have been sending signals for the response, which were blocked from producing the response by the experimental interventions. It could thus be argued that the stimulus was associated directly to a representation of the response in the central nervous system. The other paradigms provide more definitive evidence in favor of the S–S position.

Weak evidence for the S–S position is that classical conditioning still occurs when the animal is prevented from making a response.

US Devaluation Paradigm

A different test of the S–S versus S–R alternatives involves devaluing the US after classical conditioning has taken place (e.g., Holland & Rescorla, 1975; Rescorla, 1973). For instance, a CS of light can be paired with a US of food for hungry rats. The CS comes to evoke a number of responses associated with the US of food, including increased activity (Holland & Rescorla, 1975). The rats can then be satiated. When the rats are no longer hungry, will the CS (the light) still evoke activity? If the CS were directly associated to the CR, it would still produce increased activity. On the other hand, if the CS had produced increased activity through anticipation of the US of food, and the US had lost its power to produce increased activity, the CS would lose its power to evoke activity. In fact, it loses its ability to evoke activity, implying that S–S associations were formed.

This type of test is referred to as a **devaluation paradigm**. The US is devalued. If the CS is directly associated to the CR, it will not be affected by the devaluation, but if it is associated to the US, it will be affected. As another example, consider the conditioned emotional response, or CER, behaviors, such as freezing, shown by animals such as rats to CSs in anticipation of aversive stimuli, such as shock (see Chapter 1). The shock US can be devalued by repeated administration (in the absence of the CS), which makes the rat less sensitive to it (an example of habituation introduced earlier with respect to the human eye blink). When the CS is finally presented after such US devaluation, the rat shows a reduced CER. Generally, the response to the CS is reduced when the US is devalued, indicating that the CS was associated to the US.

If the US is devalued after conditioning, response to the CS is reduced, suggesting an S–S association.

Sensory Preconditioning Paradigm

Experiments using a sensory preconditioning paradigm (e.g., Rizley & Rescorla, 1972) also suggest that the CS and US become associated. A typical sensory preconditioning experiment (see Table 2.1) has two phases. In the first phase, one neutral stimulus, such as a light (CS_2), occurs just before another neutral stimulus, such as a tone (CS_1). In the second phase, CS_1 is paired with a US. For instance, the tone (CS_1) may precede a shock to the leg, which produces the UR of leg withdrawal. After a while, the tone acquires the ability to evoke leg withdrawal.

TABLE 2.1 *Three Paradigms and the Associations Formed to the CS*

Paradigm	Phase 1	Phase 2	CS_1 Association	CS_2 Association
Standard	—	CS_1–US	S–S (to US)	—
Sensory preconditioning	CS_2–CS_1	CS_1–US	S–S (to US)	S–S (to CS_1)
Second-order conditioning	CS_1–US	CS_2–CS_1	S–S (to US)	S–R or S–S (to CR or CS_1)

What happens if the other neutral stimulus, the light (CS_2), is then presented? If classical conditioning involves associating responses to stimuli, nothing will happen because the light was never paired with leg withdrawal. On the other hand, if stimuli are directly associated, the light (CS_2) will evoke the anticipation of the tone (CS_1), because of their pairing, which in turn will evoke the anticipation of the shock (US). Thus, the animal will anticipate shock in response to light and withdraw its leg. This is exactly what happens. In general, sensory preconditioning experiments show that CS_2 evokes the CR, even though it was never paired with the US and so never preceded the UR.

Sensory preconditioning shows that two neutral stimuli can be associated, consistent with the S–S position.

Second-order Conditioning Paradigm

An alternative paradigm, the second-order conditioning paradigm (see Table 2.1), provides evidence for S–R associations. Holland and Rescorla (1975) performed an experiment with rats, in which a light (CS_1) was paired with a US of food. The CS comes to evoke a CR of increased activity. Then a second stimulus, a tone (CS_2), was paired with the first, the light (CS_1). Thus, as in the preconditioning experiment, there are a CS_2–CS_1 pairing and a CS_1–US pairing. However, in the second-order conditioning paradigm, the order is reversed: first CS_1–US and then CS_2–CS_1. As in the preconditioning paradigm, the second-order CS_2 (in this example, the tone) acquires the ability to evoke the CR (in this example, increased activity). This phenomenon is not evidence for an S–S association, however. Since the first-order CS_1, the light, already evoked its own CR, it could be that this CR is now being transferred to the second-order CS_2. In effect, CS_1 could be behaving like a US and the CR like its UR. Thus, there are two possibilities: the tone could be associated to light (S–S) or increased activity (S–R).

Holland and Rescorla used a devaluation paradigm to determine whether the associations were S–S or S–R. They devalued the US of food by satiating the rats. As noted earlier, US devaluation reduces the conditioned response to the first-order CS_1, suggesting an S–S association between CS_1 and the US. However, it does not lead to reduced responding to the second-order CS_2, suggesting a direct association between CS_2 (tone) and the CR (activity). If the first-order association is extinguished by presenting light without food, the second-order association is not extinguished and the tone continues to produce increased activity (e.g., Amiro & Bitterman, 1980). Thus, the surprising result from this and other second-order conditioning experiments seems to be that first-order associations are S–S and second-order associations are S–R. However, some second-order conditioning experiments (e.g., Rashotte, Griffin, & Sisk, 1977) have found evidence for S–S second-order associations.

In second-order conditioning experiments, second-order associations tend to be S–R, in contrast to first-order associations, which are S–S.

Conclusions

The current conception (e.g., Holland, 1985a) is that both the stimulus and response aspects of subsequent events compete for association to the stimulus. Different paradigms produce S–S or S–R associations, depending on whether the subsequent stimulus or response aspects are more salient or prominent. A preconditioning experiment provides evidence for S–S associations because no salient responses are associated to the neutral CS_1 stimulus. A second-order conditioning experiment supports S–R associations because the prior first-order conditioning usually gives the CS_1 response characteristics that are more salient than its stimulus characteristics. First-order associations are more often S–S in character because the US is typically so salient. Table 2.1 attempts to summarize the results. There are three paradigms: standard, which involves one pairing, and (2) the sensory preconditioning and (3) second-order conditioning, which involve two pairings but differ in their ordering. S–S associations appear to be the rule except with respect to CS_2 associations in second-order conditioning.

Whether associations are S–S or S–R depends on whether the stimulus or response is more salient.

What Is the Conditioned Stimulus?

What exactly is the conditioned stimulus that gets associated in classical conditioning? If an organism forms an association to a tone of a particular pitch or frequency (for example, 1000 Hz), will the organism display the association to a slightly different pitch (1010 Hz), a very different pitch (4000 Hz), an entirely different sound (a dog barking), or a flash of light? Intuitively, it would seem that a slightly different stimulus, such as 1010 Hz, should elicit the association but that a very different one (the flash of light) should not.

Siegel, Hearst, George, and O'Neil (1968) conditioned the eyelid response in rabbits to tones of different frequencies. Different rabbits experienced tones with frequencies of 500, 1000, 2000, 3000, or 4000 Hz paired with a puff of air to the eye. They were then tested with a variety of stimuli at different frequencies. Figure 2.8 shows the **generalization gradients** obtained for each training stimulus. The amount of eyelid response to various sounds falls off as the sounds become more different from the original training stimulus. For instance, a rabbit trained with 1000 Hz responded most often to a test of 1000 Hz, next most frequently to 500 or 2000 Hz, then to 3000 Hz, and least frequently to 4000 Hz.

Subjects can learn to restrict the range of stimuli to which they respond. This phenomenon is called **discrimination learning**. In one experiment with humans, Gynther (1957) trained an eyeblink response to a CS involving a dim light to the right. After conditioning, when initially tested with a dim light to the left, his subjects showed a strong tendency to respond as well. However, the US (puff of air to the eye) was never associated with the light to the left. After 100 discrimination trials (50 with each CS), subjects discriminated between the two stimuli and showed a much greater tendency to blink to the right light.

Experiments such as these show that classical conditioning can generalize to a range of stimuli and that this range of stimuli can change as the subject learns whether or not the CSs are associated with the US. The phenomena of stimulus generalization and discrimination are discussed at greater length in the next chapter inasmuch as they have received more attention in research on instrumental conditioning.

Organisms naturally generalize a CR to a range of similar CSs, but they can be trained to change the range of CSs to which they emit the response.

What Is the Conditioned Response?

Two possibilities exist as to the nature of the CR in classical conditioning. First, the CR may be a version of the UR. Perhaps the CS causes the organism to experience some internal image of the US, which then produces the same behavior for whatever reason it was produced by the US. The other view is that the CS is informational; it allows the organism to anticipate that the US will occur and thus to take appropriate action in anticipation of the US. In this view, the CR is a preparation for the US, not a response to it. It is noteworthy that the CRs that are conditioned are adaptive; that is, the CRs tend to prepare the organisms for the US. As noted earlier in this chapter, it is to the human's benefit to blink in anticipation of a puff of air. It is also to the dog's benefit to salivate if food is coming and to flex its leg in anticipation of shock. The original US–UR reflex is

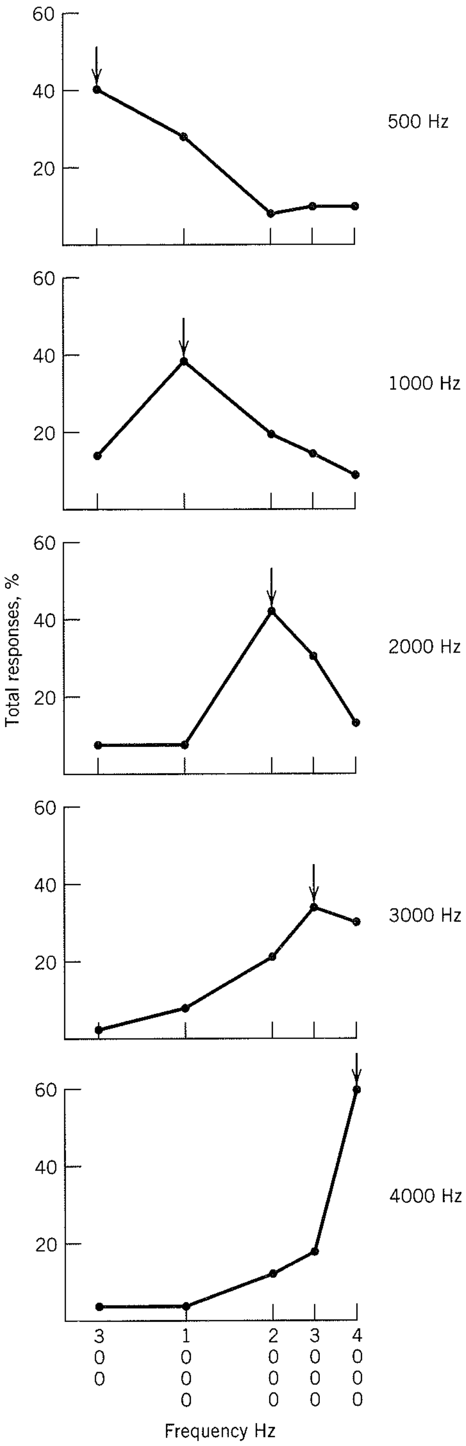

FIGURE 2.8 Mean percentage of total generalization test responses for groups trained at each CS value. (The arrow indicates the frequency of the training stimulus for each group.) *Source:* From S. Siegel, E. Hearst, N. George, and E. O'Neal. Generalization gradients obtained from individual subjects following classical conditioning. *Journal of Experimental Psychology,* Volume 78. Copyright © 1968 by the American Psychological Association. Reprinted by permission.

in the organism to begin with because it is adaptive. Thus, the CS–CR connection also tends to be adaptive.

One kind of evidence for the preparatory character of the CR comes from cases in which the CR is not the same as the UR. A striking example concerns heart-rate changes in response to electric shock. The unconditioned response to electric shock is heart-rate acceleration; but in some organisms the conditioned response is heart-rate deceleration (Obrist, Sutterer, & Howard, 1972; Schneiderman, 1973). In anticipation of a shock, the organism relaxes, which may reduce the perceived magnitude of the pain. This relaxing reduces the heart rate. Similarly, some organisms, such as rats, tend to show increased activity in response to shock but freeze in response to a stimulus associated with shock (Rescorla, 1988b). Thus, the behavioral responses to the CS can be quite the opposite of the responses to the US, but both can be viewed as adaptive responses. The shock is a noxious stimulus that requires escape, whereas the CS is treated as a warning that may require a different response. For example, in nature when an animal sees a predator (CS), freezing may help it avoid detection, but once the predator attacks (US), flight is the appropriate response.

A number of unconditioned responses involve a biphasic structure in which an initial response is followed by an opposing response. One of the most dramatic involves responses to narcotics, such as heroin. The initial response to heroin, a feeling of euphoria, is followed by a second, **opponent response**, which tends to counteract the initial response and produces unpleasant withdrawal symptoms. Another example of such an opponent process is skydiving (Epstein, 1967). The feelings of terror the skydiver has just before the dive are followed by an antagonistic response of pleasure when the dive is successful. Solomon and Corbit (1974) suggest that the antagonistic opponent response is caused by the body trying to avoid extreme arousal states, which are demanding of resources. Also, in the case of heroin, the increased pain threshold it produces can also be dangerous.

Wagner (1981) proposed that in the case of such biphasic URs, the compensatory second response is conditioned to the CS because it is an appropriate response to blunt the effect of the US. On the other hand, when the UR is monophasic, like an eye blink, and does not involve an opponent process, the UR is conditioned because the UR is the appropriate response in anticipation of such a US.[2] Wagner called his theory **SOP**, for **sometimes opponent process**, because sometimes (i.e., in the case of biphasic responses) the conditioned response is the opponent process.

It has been suggested that the conditioning of the opponent process is responsible for drug tolerance (Siegel, 1983). With repeated use of a drug, the high becomes weaker and weaker; consequently, the addict needs to take more and more of the drug to produce the same high. It is claimed that the stimuli that accompany drug use (for example, a needle) serve as a CS that becomes conditioned to a second, opponent process. Thus, the conditioning of the opponent process is producing the drug tolerance. The CR evoked by these stimuli is the opponent process that negates the effect of the drug. There is evidence that much of the tolerance built up to drugs such as morphine can be removed by changing the stimuli that accompany their administration (for a review, see Siegel, 1983).

---

[2] However, even in the case of the eye blink, Wagner distinguished two components and argued that only the second component becomes conditioned. His position is supported by Thompson's research on the circuitry underlying eyeblink conditioning reviewed earlier. This research showed that the unconditioned eye blink had two components: one with a latency of 20 msec and one with a latency of 70 msec. The longer-latency component forms the basis for the CR.

Akins, Domjan, and Guitierez (1994) provide an interesting example of how the CR adapts to the circumstances of the US. They were looking at conditioning of sexual behavior in male quail where the US was access to a female and the CS was the appearance of foam block with bright orange feathers where the female would eventually appear. This CS evoked a CR of searching behavior, which is different from the UR to the US, which would be to engage in courtship and copulatory responses. Again we can see that the CR is behavior in anticipation of the US. In this case, the quail is looking for the female that the CS signals.

In addition to this basic result, Atkins et al. manipulated the delay between the CS and the US in the conditioning experiment. They used either 1 minute or 20 minutes. These are long CS–US intervals compared to the experiments we have considered so far (e.g., Figure 2.3). As we will see later in the chapter, however, conditioning can sometimes be obtained at even longer CS–US intervals than 20 minutes for some kinds of US. Atkins et al. were interested in the difference in the nature of the conditioned searching behavior at these two delays. They classified the searching behavior as focused (the quail actually approached the CS) or general (the quail just ran around). Figure 2.9 compares the percentage of general versus focused search in the two conditions. As can be seen,

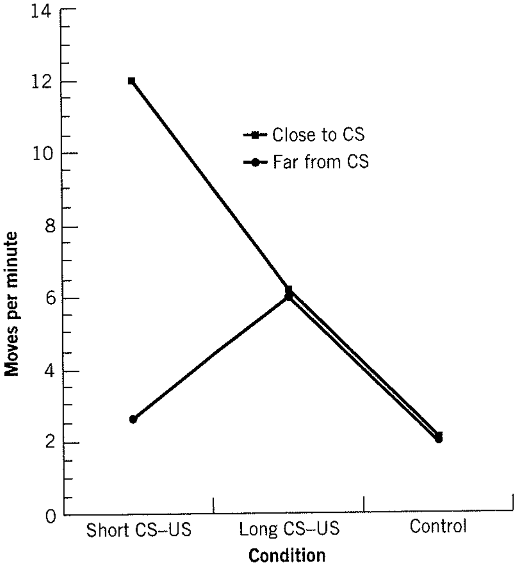

FIGURE 2.9 Amount of movement close to the CS and far from the CS as a function of the duration of the CS–US interval.

search focused near the CS predominated when the CS–US interval was 1 minute, whereas there was general search everywhere (equal amount of movement close and distant from goal) when it was 20 minutes. Again, this makes sense. If the CS signals a mate in 20 minutes, then the female can be anywhere and the quail needs to find her, whereas if the mate is about to appear in a minute, the quail should approach the CS and wait for his impending opportunity. Thus, not only is the CR an adaptation to the US, but the nature of the adaptation can vary appropriately with things like the CS–US interval.

The CR (which is not always the same as the UR) is often an adaptive response in anticipation of the US.

Association: The Role of Contingency

An idea that extends back at least to Aristotle's writings on associations is that two things become associated when they occur together in time and space. This is the principle of contiguity. Some associative theories have claimed that **contiguity** of CS and US is sufficient for classical conditioning. Earlier in this century, contiguity was thought to be critical for conditioning. Evidence for its importance was the strong influence that the CS–US interval can have on conditioning. For instance, Figure 2.3 showed that eyeblink conditioning was maximal if the CS occurred very close in time to the US. However, as we have just seen with respect to Figure 2.9, in some of the more recent research, conditioning has been obtained at long delays. Thus, the importance of contiguity has been downplayed in recent theories of conditioning.

An alternative view is that conditioning occurs only if there is a **contingency**; that is, the first stimulus predicts the occurrence of the second. In this view, mere co-occurrence of CS and US is not sufficient to form an association. For instance, when my two sons watch TV, they argue. However, my two sons argue a lot when they are involved in other activities, too. Thus, the mere co-occurrence of TV and arguments does not mean that there is a contingency (i.e., watching TV causes them to argue). The probability of arguments must be greater when the TV is on for there to be a contingency. Stated symbolically, a predictive or contingent relationship requires

$$P(\text{argument} \mid \text{TV}) > P(\text{argument} \mid \overline{\text{TV}})$$

where $P(\text{argument} \mid \text{TV})$ is the probability of an argument when the TV is on and $P(\text{argument} \mid \overline{\text{TV}})$ is the probability of an argument when the TV is not on.

Rescorla's Experiment

Rescorla (1968b) conducted an experiment to determine whether contiguity or contingency is essential in classical conditioning. He intermittently presented a 2-min tone while rats were pressing a bar. In different conditions, he presented

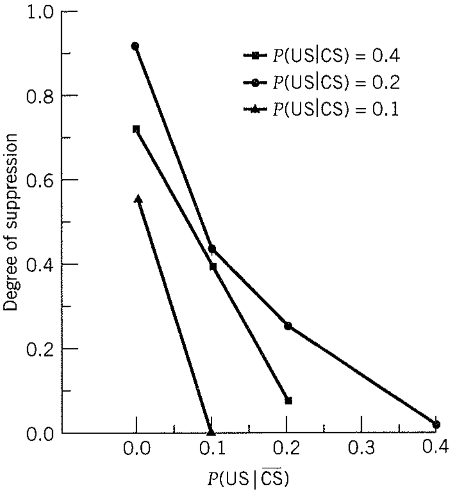

FIGURE 2.10 Dependence of conditioning on both $P(\mathrm{US}\,|\,\mathrm{CS})$ (likelihood of the US during the CS interval) and $P(\mathrm{US}\,|\,\mathrm{CS})$ (likelihood of the US in the absence of the CS). (Adapted from Rescorla, 1968.)

shock during 10, 20, and 40 percent of these 2-min tone intervals. Rescorla was interested in the degree to which the rats would show a CER (freezing) and decrease their rate of bar pressing during a tone interval.

Rescorla also varied the probability of shock during the 2-min intervals when no tone was present, creating three separate conditions with 10 percent, 20 percent, or 40 percent probability of a shock during the no-tone intervals. Figure 2.10 shows the results for various combinations of probability of shock in the two types of intervals. The plot shows how much less the rats pressed the bar in the presence of a tone relative to how much they pressed in the absence of a tone. If they pressed P times in the presence of a tone and A times in the absence of a tone, the measure of suppression is the proportion $(A - P)/(A + P)$.[3] Since P and A are measured over the same periods of time, they should be equal if the rat is unaffected by the tone, and the suppression rate should be zero. To the extent the animal is conditioned to the tone, P should be near zero, and the suppression rate should be near one, indicating that the animal is freezing in the presence of the tone.

Consider the degree of conditioning displayed when the probability of a shock, $P(\mathrm{US}|\mathrm{CS})$, was .4 during a tone interval. The amount of conditioning is a function of the probability of a shock when no tone is present, $P(\mathrm{US}|\overline{\mathrm{CS}})$. When the probability of a shock in a no-tone interval was zero, the classical conditioning result of large response suppression occurred. The degree of suppression was near one, indicating that the rats almost never pressed the bar in the presence of the tone. However, as the probability of a shock in the no-tone interval increased to .4, the level of suppression decreased. When the probability of shock was the same in the presence or absence of a tone (i.e., both were .4), no conditioning occurred; the degree of response suppression was zero, implying that the rate of bar pressing was the same in both the presence and absence of the tone. The other conditions in Figure 2.10 displayed similar results. The amount of conditioning is not a function of contiguity (frequency with which shock and tone co-occur) but rather of contingency (difference between probability of shock given tone and probability of shock given no tone).

---

[3] This is an algebraic transformation of the measure reported by Rescorla.

Conditioning only occurs when presentation of the CS is associated with increased probability of the US.

Conditioned Inhibition

The previous paradigm demonstrated that organisms show conditioning when the probability of the US is greater in the presence of the CS. What if the probability is lower? In that case there is evidence that the organism acts as if it expects the US not to occur. This phenomenon is called **conditioned inhibition**.

The standard paradigm for demonstrating conditioned inhibition involves using two CSs: a CS+, which is positively associated to the US, and a CS–, which is negatively associated to the US. In an experiment by Zimmer-Hart and Rescorla (1974), when a clicker (CS+) was presented, a shock followed, but when both a clicker (CS+) and a flashing light (CS–) were presented, no shock occurred. (The CS– was never presented alone during this initial training.) The animal came to show conditioning of the CER to the CS+ alone but not to the combination of CS+ and CS–.

How does the animal treat the CS– in this paradigm? There are a number of ways to show that CS– leads to anticipation of no shock. One method, the summation test (originally devised by Pavlov), looks at what happens when the CS– is presented with another CS that has been associated with the US. So, for instance, Zimmer-Hart and Rescorla also conditioned a 1200-Hz tone to shock. When the tone and flashing light (CS–) were presented together, the CER was reduced. Thus, the CS– negates anticipation of the US to any CS. A second paradigm, the retardation paradigm, involves trying to condition the CS– to the US. It takes longer to condition the CS– than to condition a new CS, indicating that the animal has to overcome an association of CS– with the anticipation of no US.

Sometimes the response that is being conditioned is bidirectional, and the organism can indicate conditioned inhibition by doing the opposite of the CR. For instance, Wasserman, Franklin, and Hearst (1974) trained pigeons to associate food with an experimental context (CS+). A light (CS–) signaled that there would not be food and so became a conditioned inhibitor. In the experimental context (CS+), the animals approached the feeding apparatus, but in the presence of the light they actively withdrew from the light. Another example involves taste preference. Association of taste with illness decreases the preference for the taste, but the association of the flavor with the absence of illness actually increases the preference for that flavor (Best, Dunn, Batson, Meachum, & Nash, 1985).

Organisms can be trained to associate a CS with the absence of a US, and the CS then becomes a conditioned inhibitor for responding.

Associative Bias

This chapter has reviewed the evidence that organisms are sensitive to the statistical regularity between a CS and a US. There is also evidence, however, that they are predisposed to associate certain CSs to certain USs independent of their statistical regularity. This preference for certain associations is referred to as **associative bias**, which is similar to Thorndike's concept of belongingness discussed in the first chapter.

Suppose you heard a loud noise that sounded like an explosion followed by the shaking of the ground. How likely would you be to think the explosion sound and the shaking of the ground were related? Suppose, on the other hand, you heard a bird's song followed by the shaking of the ground? How likely would you be to think that the bird sound and the shaking were related? Presumably, you would think the first pair was more probably related than the second pair. Certain pairs of stimuli are more likely to be related than are other pairs of stimuli, and organisms condition more readily with such pairs of stimuli.

In a well-known study, Garcia and Koelling (1966) used rats whose only access to water was from a water spout. While drinking the water, the rats were exposed to a compound CS consisting of a flavor component (a saccharine taste) and an audiovisual component (light flash and click). Different groups of rats then received either a shock or an injection of a drug that produced nausea. Garcia and Koelling tested the rats separately with saccharin-flavored water and with the audiovisual stimulus to see which had a greater impact on water intake. Rats that had received a shock later showed greater CER (i.e., reduced drinking) to the audiovisual stimulus, whereas rats that had received the injection showed greater CER to the flavor stimulus. Thus, rats were more prepared to associate a CS of light with a US of shock and a CS of flavor with a US of poisoning. A great deal of subsequent research has focused on conditioned taste aversions of a distinctive taste (CS) with a poisoning (US). Rats (and many other organisms, including humans) learn such taste aversions after a single CS–US pairing and after intervals between the CS and US up to 24 hours (Etscorn & Stephens, 1973). It is unusual for conditioning to operate over such long time delays (e.g., see Figure 2.3). Organisms presumably have this unusually strong propensity to associate taste and poisoning because it is adaptive.

The associations that organisms are prepared to form are somewhat species specific. Wilcoxon, Dragoin, and Kral (1971) compared using the CS of water taste (sour) or water color (dark blue) in conditioning to a US of poisoning . They compared rats and quail and found that rats displayed greater aversion to taste and quail displayed greater aversion to color. This result makes sense when the two species are compared. Rats are nocturnal animals with an excellent sense of taste and smell but with poor vision; quail are daytime animals with excellent vision, and they use this vision to select appropriate foods (e.g., seeds that are not poisonous).

In some situations, a conditioned food aversion can be a serious problem. For instance, cancer patients undergoing chemotherapy develop strong aversions to the foods they eat because chemotherapy produces nausea; thus, they do not eat adequate food. One means of combating this problem is to schedule eating and chemotherapy so that food intake does not precede nausea. Another possibility for patients is to consume a bland diet, because nondistinctive stimuli do not appear to become conditioned as readily as distinctive stimuli. Yet another possibility is to have patients consume a very distinctive-tasting food before the onset of nausea so that the taste aversion will only be conditioned to that stimulus. For example, the problem of taste aversion with children was reduced if they received a distinctive-tasting "mapletoff" ice cream before chemotherapy (Bernstein, Webster, & Bernstein, 1982).

Organisms have a bias for associating stimuli that are likely to be related in their environment.

Conclusions about the Nature of the Association

Associations are formed in the classical conditioning paradigm when the appearance of the CS increases the probability that the US will occur. When this is not the case, associations are not formed even if there is a high degree of co-occurrence between the CS and the US. An organism is also sensitive to the likelihood of an association between the two stimuli (associative bias). The organism might be characterized as forming a statistical inference. Bayesian statistical models for inference about probabilistic relationships consider not only the data but also prior beliefs about what relationships are likely. The fundamental relationship in Bayesian statistics is[4]

$$\text{Posterior belief} = \text{Prior belief} * \text{Evidence}$$

That is, the posterior, or final, belief in a hypothesis (e.g., that the CS predicts the US) is a product of the prior belief and the strength of the evidence for the

---

[4] For those who prefer the odds formula in terms of probabilities:

$$\frac{P(H \mid E)}{P(\overline{H} \mid E)} = \frac{P(H)}{P(\overline{H})} * \frac{P(E \mid H)}{P(E \mid \overline{H})}$$

where $P(H \mid E)$ is the posterior probability of the hypothesis given the evidence; $P(\overline{H} \mid E)$ is $1 - (P \mid H)$; $P(H)$ is the prior probability of the hypothesis; $P(\overline{H})$ is $1 - P(H)$; $P(E \mid H)$ is the conditional probability of the evidence given the hypothesis; and $P(E \mid \overline{H})$ is the conditional probability of the evidence given that the hypothesis is false.

hypothesis. Organisms in a conditioning experiment can be viewed as acting according to this statistical prescription. Posterior belief would map onto strength of conditioning between the CS and the US, prior belief onto a bias to associate the CS and the US, and evidence onto the degree of contingency between the CS and the US.

> *Classical conditioning may be viewed as forming associations that are well-founded inferences from a statistical perspective.*

Conditioning to Stimulus Combinations

The research just reviewed shows that organisms respond to the predictiveness of individual stimuli. This sensitivity to the predictive structure of the stimulus situation has been further documented by a number of lines of research studying conditioning when multiple stimuli are present as part of the CS.

Blocking

Experiments have shown that an association is not formed with one CS if another CS is more informative. In one experiment, Kamin (1968) contrasted two groups:

> **Control.** The animals experienced eight trials in which CSs of noise and light were followed by shock.
>
> **Experimental.** The animals received 16 trials in which just the noise was followed by the shock. Then, like the control condition, the animals received eight trials in which a CS of noise and light was followed by shock.

Kamin conducted separate tests to determine whether the CER could be evoked to noise or to light. He found that the CER had been conditioned to both noise and light in the control group but only to noise in the experimental group. Thus, for the experimental group, the CER was not conditioned to the light even though there were eight reliable pairings of light and shock, which normally would have produced conditioning. The noise had already been established as a reliable predictor of the shock, and the light added no more information; therefore, no association was formed. The tendency for one stimulus to overshadow another stimulus is called **blocking**.

In a variation of this paradigm, Kamin (1969) presented a more intense shock to the light–tone combination than to the tone alone. The animals were first given 16 trials of tone–shock pairings with a shock intensity of 1 mA (milliampere), followed by eight trials of tone and light followed by shock with an intensity of 4 mA. In this condition, rats showed significant conditioning of the

CER to the light. The light was now informative because it signaled a more intense shock.

In another variation of this paradigm, Wagner (1969) trained three groups of rabbits in an eyeblink-conditioning paradigm in which the US was a 4.5-mA shock to the area of the eye:

Group 1. Two hundred trials of light and tone followed by shock.

Group 2. Two hundred trials of light and tone followed by shock, intermixed with 200 trials of light followed by shock.

Group 3. Two hundred trials of light and tone followed by shock, intermixed with 200 trials of light followed by no shock.

In contrast to Group 1, which defines the reference condition, Group 2 showed little conditioning of the eyelid response to the tone. As in Kamin's studies, this result can be interpreted to mean that the light was the better predictor. On the other hand, Group 3 showed even greater conditioning of the tone to the eyelid response than did the reference Group 1. The rabbits in this group were getting evidence that the light was not a good predictor of shock and so came to treat the tone as the sole reliable predictor of shock.

When one CS is a more reliable predictor of a US than a second CS, that CS blocks conditioning of the other CS.

Configural Cues

In the experiments discussed thus far, separate stimuli developed separate associations to the US. However, it is possible to condition an organism to respond only if a particular configuration of stimuli are present. Organisms can be trained to respond when both stimuli A and B are present and not when just one is present. Although this result could mean that the AB combination is associated, it could also indicate that A and B are separately associated to the US but are too weakly associated to evoke the CR individually and are only strong enough to do so in combination. Whereas it is possible to explain this result in terms of separate associations, in some situations the only possible explanation is that a configuration of stimuli have become associated to the response and that the individual stimuli are not associated separately. For example, organisms can also be trained to respond when A is present and when B is present, but not when AB is present (see Kehoe & Gormezano, 1980, for a review). If there is a positive strength association between A and the US and between B and the US, an even stronger association would be expected when both A and B are present. However, it appears that associations can be learned to cue combinations. In this case, the cues A + noB and noA + B become associated to the US, but the combination A + B does not. Another situation that shows configural associations involves four cues—A, B, C, and D. Organisms can be taught to associate

A + B and C + D to the US but not A + C or B + D. Again, it is not possible to account for this result in terms of associations to the individual stimuli. Associations must be formed to the combinations. Such stimulus combinations are referred to as **configural cues**.

This phenomenon is quite different from blocking, where one cue (A) in a combination masks the other (B). In the typical blocking paradigm, the organism comes to respond to A alone and to AB in combination, but not to B alone. In a typical configural conditioning experiment, the organism comes to respond to A alone to B alone, but not to AB (although as described earlier there are other ways to demonstrate configural conditioning). Configural conditioning also differs from a conditioned inhibition paradigm, where one cue negates the effect of the other and acquires the properties of a conditioned inhibitor (the organism learns to respond to A alone, but not to AB or to B alone). In a typical configural conditioning paradigm, both cues maintain positive effects but only when presented alone.

Configural conditioning involves learning associations to combinations of stimuli that are different from the associations to the individual stimuli.

Conclusions

In some cases (blocking and conditioned inhibition), it seems that the response to the combination of stimuli can be predicted from the response to the individual stimuli, but in other cases (configural cues) it cannot. The next section reviews the Rescorla–Wagner theory, which accounts for many of the occasions when the response to the stimulus combination can be understood in terms of the response to the individual stimuli.

The Rescorla–Wagner Theory

In 1972, Rescorla and Wagner proposed a theory that successfully predicts many phenomena of classical conditioning. Their theory shows how simple learning mechanisms can be sensitive to the contingency between the CS and the US. It also illustrates the principle that simple mechanisms can produce the highly adaptive statistical sensitivities documented in the previous section. Although the theory is 25 years old, it has gained renewed currency as a popular theory of neural learning.

The **Rescorla–Wagner theory** has three basic constructs. One construct is the strength of association, V, between CS and US that controls the amount of conditioning displayed. The second is the maximum level, λ, that this strength can reach, thought to be a function of the US such that stronger USs can maintain stronger associations. The third construct is the rate of learning, α, that determines how rapidly this limit can be reached. The fundamental assertion in the theory is that when a CS and a US are paired, the strength of their association increases proportionally to

the rate of learning, α, and

the difference between the maximum strength possible and the current strength, $(\lambda - V)$.

Cast as an equation, this becomes

$$\Delta V = \alpha\,(\lambda - V)$$

where ΔV is the change in strength. One can view λ as reflecting the strength of the actual US and V as reflecting how strong the US was predicted to be. The theory states that there will be learning to the degree that there is a difference between what was predicted and what occurred. It is sometimes said that the Rescorla–Wagner theory predicts learning to the degree that the US is surprising (i.e., not predicted by the CS).

Consider the application of this theory to a simple situation in which the US and the CS are paired 20 times. Assume that the maximum strength of association, λ, is 100 and that the rate of learning, α is .20. The initial strength of association is zero. After the first learning trial, the increase in the strength of association is

$$\Delta V = .20\,(100 - 0) = 20$$

The strength of association is the sum of this increment plus the prior strength of 0: $20 + 0 = 20$. In the next trial, the same formula applies, except that the prior strength is now 20 instead of zero, and so the increment is

$$\Delta V = .20\,(100 - 20) = 16$$

The total strength is this result plus prior strength: $16 + 20 = 36$. This process can be continued, calculating for each trial the total amount of strength. Figure 2.11 shows the growth of strength over the first 20 trials; the growth looks like a typical conditioning function (e.g., see Figures 1.4 and 2.2).

> ***The Rescorla–Wagner theory asserts that learning is proportional to the difference between the current strength of association and the maximum strength that the US will allow.***

Application to Compound Stimuli

The Rescorla–Wagner theory is similar to several mathematical theories developed to account for learning, and it is not surprising that it succeeds in accounting for the approximate form of the conditioning curve. The theory derives spe-

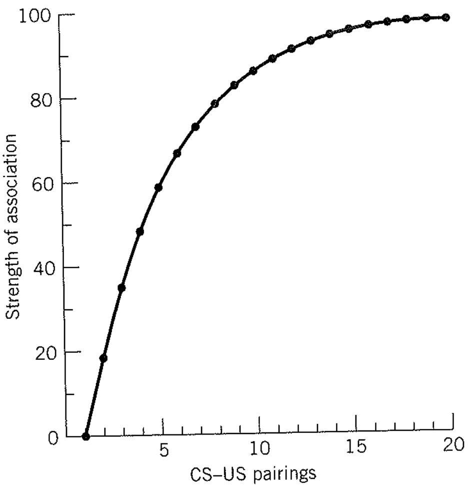

FIGURE 2.11 Growth of associative strength with repeated CS–US pairings, according to Rescorla–Wagner theory.

cial interest from how it deals with compound cues such as those used in blocking experiments. Before considering blocking experiments, consider a simpler situation. Suppose that two stimuli, A and B, are simultaneously presented as the CSs for the US. For example, the US might be food, A might be a tone, and B might be a light. The Rescorla–Wagner theory holds that the total strength of association between the compound cue and the US, which can be denoted V_{AB}, is the sum of the strengths of the individual associations of A and B to the US, which can be denoted V_A and V_B. That is,

$$V_{AB} = V_A + V_B$$

When A and B are paired with the US, they will grow in strength of association to the US as in the case where there is only one CS:

$$\Delta V_A = \alpha(\lambda - V_{AB})$$

$$\Delta V_B = \alpha(\lambda - V_{AB})$$

On the first trial, since there is no prior conditioning, these equations become (assuming $\alpha = .20$ and $\lambda = 100$):

$$\Delta V_A = .20(100 - 0) = 20$$

$$\Delta V_B = .20(100 - 0) = 20$$

Thus after the first trial, the individual stimuli have strengths of 20 and V_{AB} is the sum, or 40. On the next trial, the equations become

$$\Delta V_A = .20(100 - 40) = 12$$

$$\Delta V_B = .20(100 - 40) = 12$$

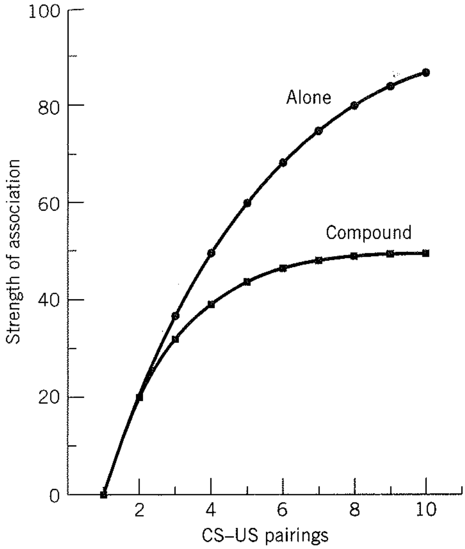

FIGURE 2.12 Comparison of growth in strength to cue alone versus cue in compound combination.

and the individual stimuli have strengths 20 + 12 = 32. Figure 2.12 shows the growth in the strength of association to a single cue when the cue is part of a compound versus when it is alone. The cue strength of A can only reach 50 because it must share the strength of association with B, which will get the other 50. This example illustrates an important feature of the Rescorla–Wagner theory—that the various stimuli compete for association to the US, a situation often referred to as **competitive learning**.

In the preceding equations, α is the same for A and B. But Rescorla and Wagner allow for the possibility that the stimuli may vary in their salience and therefore one stimulus may have a more rapid learning rate. Suppose that the rate of learning for A is .4 and for B it is .1. Figure 2.13 plots the rate of strength accumulation for the two stimuli. As shown, A overshadows B and acquires most of the strength. When one stimulus is sufficiently less salient than another, it fails to be conditioned when presented in a compound cue, even though it can be conditioned when presented alone (e.g., Kamin, 1969).

When there are multiple CSs, the Rescorla–Wagner theory assumes that they will compete for the total associative strength.

Application to Blocking and Conditioned Inhibition

The Rescorla–Wagner theory can be applied to the phenomenon of blocking. Recall the original Kamin experiment in which rats were given an initial 16 trials pairing A with the US followed by eight trials pairing A and B with the US.

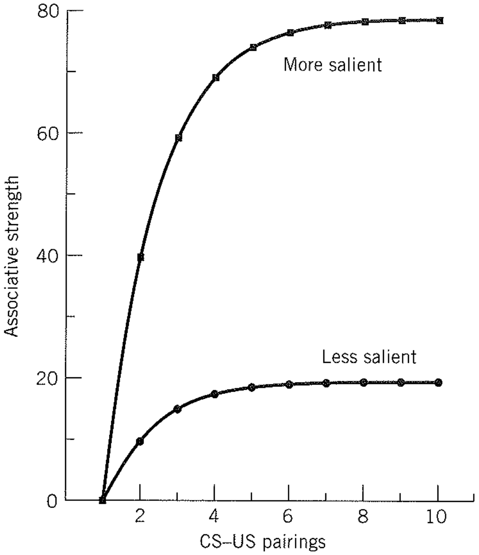

FIGURE 2.13 Growth in strength of association to a more versus a less salient stimulus.

The effect of the first 16 trials is that A has already acquired most of the available strength and nothing is left for B. That is, the effect of the conditioning is to set the strength of association to A to full value.

$$V_A = 100$$

V_B will start with an associative strength of 0, but the stimulus combination will have a strength of 100. That is,

$$V_A + V_B = 100$$

Since the US strength λ is also 100, there will be no difference between it and the strength of the compound stimuli. Since there is no difference between the US and the strength, there will be no learning and V_B will stay at 0.

The Rescorla–Wagner theory also explains why B conditions if a stronger shock is used for the AB combination than for A alone. Rescorla and Wagner postulated that the value of λ was related to the intensity of the US. Thus, if a higher value is used for the shock, as in Kamin (1969), the value of λ is larger and strength is available to be conditioned to B. As noted earlier, B can be conditioned if the shock intensity is increased.

The Rescorla–Wagner theory can also predict the phenomenon of conditioned inhibition discussed earlier. In the paradigm, A (e.g., a clicker) is associated with the US (e.g., a shock) but AB (e.g., clicker plus flashing light) is not. The Rescorla–Wagner theory implies that the organism will learn strengths of association, V_A to the clicker and V_B to the light, such that

$$V_A = 100$$

$$V_A + V_B = 0$$

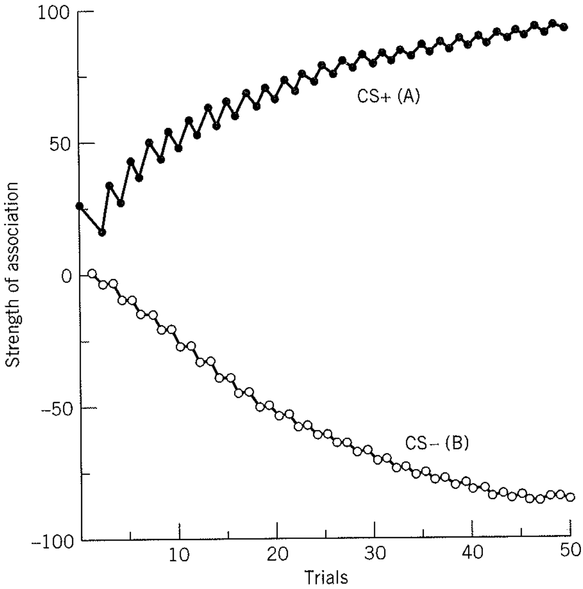

FIGURE 2.14 Growth in strength of association for the positive and negative CSs in a conditioned inhibition paradigm. *Source:* W. Schneider, "Behavior Research Methods, Instruments, and Computers," "Connectionism: Is it a paradigm shift for psychology?" (1987). Reprinted by permission of W. Schneider.

since A is always associated with shock but the AB combination never is. The only way these equations can hold is for $V_B = -100$. Thus, B has a negative strength of association that corresponds to the result of conditioned inhibition. Figure 2.14 shows the growth in strength to the positive CS+ (e.g., clicker) and the negative CS– (e.g., light) over trials, assuming that the trials presented to the organism alternate between A (CS+) and AB (CS+ and CS–). The learning curves go to +100 and –100.

The Rescorla–Wagner theory predicts blocking and conditioned inhibition because of its assumption that stimuli compete for association to the US.

Problems with the Rescorla–Wagner Theory

The Rescorla–Wagner theory explains a wide variety of experimental data, but there are some things it does not explain. A phenomenon that proves difficult for the theory concerns the effect of stimulus preexposure. If an organism is exposed to the CS a number of times before an experiment begins and then is given CS–US pairings in the experiment, the rate of conditioning is impaired (Reiss &

Wagner, 1972). The phenomenon of preexposure slowing conditioning to a stimulus is called **latent inhibition**. The organism has been exposed to a number of occurrences without anything happening and has a bias to maintain a belief in CS ineffectiveness after the conditioning trials have begun. However, there is no place in the Rescorla–Wagner theory for such an effect. The CS–US strength starts out at zero, and with no US presentations in the preexposure trials, it stays at zero. Thus, the Rescorla–Wagner theory predicts that conditioning should proceed as if there had been no preexposure trials. It has been suggested that the rate of learning, α, should reflect the salience of the CS and that the effect of CS preexposure should be to make it less salient. Wagner (1978) proposed an extension of the Rescorla–Wagner theory that includes, in effect, this assumption.

The Rescorla–Wagner theory focuses on the US in explaining learning and claims there is learning as a function of how surprising the US is. Thus, it cannot explain manipulations like latent inhibition which only involve manipulations of the CS. Therefore, other theorists (e.g. Mackintosh, 1975; Pearce & Hall, 1980) have focused on the role of the CS in conditioning. Their proposal is that the organism will pay attention to the CS and learn CS–US relationships to the extent that the CS is followed by unexpected US events. In the case of latent inhibition, the animal has had a history of the CS not being followed by significant US events and has learned to ignore the CS. Thus, in these theories the competition occurs, not between different associations to a single US, but among potential CSs for attention of the organism.

Among the triumphs of the Rescorla–Wagner theory is its treatment of blocking in which the two stimuli combine independently. However, the research on configural cues indicates that the stimuli are not always independent and stimulus combinations can be conditioned. Such results can only be accommodated within the Rescorla–Wagner theory by introducing compound stimuli as CSs different from the individual stimuli. Thus, not only are A and B treated as stimuli that can be conditioned, but so is the compound AB. This idea was proposed by Spence in 1952 and suggested by Rescorla and Wagner, and it has been actively developed by Gluck and Bower (1988) in an application of the Rescorla–Wagner theory to human learning. However, the introduction of compound stimuli weakens the predictive power of the Rescorla–Wagner theory because there is no basis for predicting whether the cues will be treated separately or configurally.

The data on associative bias also present problems for the Rescorla–Wagner theory because they indicate that learning is not just a function of the CS or the US, but depends on the interaction between the two. Some pairs belong together and are easier to associate; for example, rats are especially prepared to condition a US of poisoning to a CS of taste. The Rescorla–Wagner theory can accommodate this phenomenon by assuming that the learning rate, α, varies with the CS–US combination. As in the case of compound cues, however, such a maneuver weakens the theory because the theory offers no basis for knowing how to assign α's to CS–US combinations.

As we have observed, the Rescorla–Wagner theory is not without its problems, and alternative theories of conditioning have been proposed. We have already mentioned the theories that focus on the attention the organism pays to the CS. Yet other theories propose that the organism compares the relative effectiveness of various CSs in predicting the US and shows conditioning to the extent a CS is more effective than the alternatives (Gibbon & Balsam, 1981; Miller & Matzel, 1989). However, these theories have their own difficulties. Suffice it to say that no theory can capture the many phenomena that have been documented with classical conditioning.

We have focused on the Rescorla–Wagner theory because it has remained the reference theory for the field for more than a quarter of a century. In 1997 at the meeting of the Psychonomics Society, a symposium was held to commemorate the twenty-fifth anniversary of that theory. Few theories survive that long in psychology, and at that symposium Rescorla offered some speculations to explain the durability of the theory. He proposed that one of its advantages was that it was relatively simple and captured many of the phenomena of interest to the field. It helped set the research agenda for the next 25 years as people explored the strengths and weaknesses of the theory and developed alternatives. It captured the basic idea that conditioning involved learning to predict the US and that there would be learning to the degree that the US was not predicted and so was surprising. The basic learning rule has been used by a number of neural models. At the same symposium, Gluck proposed that the Rescorla–Wagner learning rule described one kind of low-level neural learning but that often other kinds of neural learning were taking place as well. In the next section we consider how the Rescorla–Wagner rule has been used as a model of neural learning.

The Rescorla–Wagner theory captures some of the ways organisms are more sensitive to the statistical relationships among stimuli, but organisms show more sensitivity than the theory can capture.

Neural Realization: The Delta Rule

The Rescorla–Wagner theory corresponds to a current idea about how learning may take place at the neural level, which has played a particularly important role in a theory of neural processing called **connectionism**. This theory stresses the importance of synaptic connections among neurons. Figure 2.15 illustrates a typical connectionist processing module where neural activation comes in along a set of input neurons (at the bottom), each of which synapses onto a set of output neurons. Every input neuron is associated to every output neuron. The pattern of connections presents a more complex situation than that found in *Aplysia* (Figure 2.5), where one neuron (a motor neuron) becomes more active when another neuron (a sensory neuron) becomes active.

Researchers have maintained that the greater complexity of such networks is required in order to reflect the greater complexity of mammalian behavior and

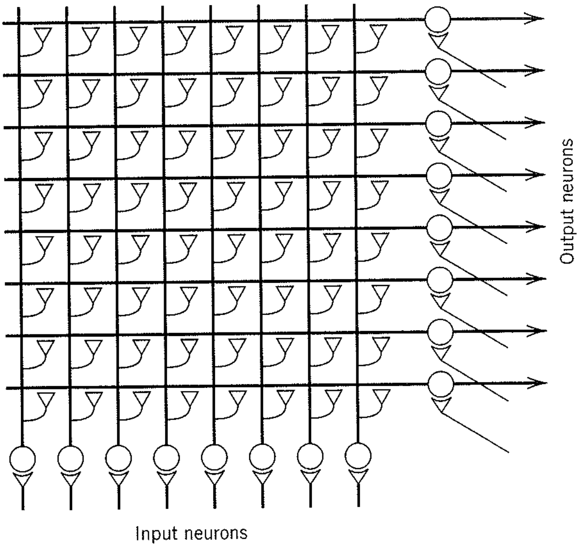

FIGURE 2.15 Schematic representation of a neural net, where input neurons come in (at the bottom) and synapse on output neurons (at the right). All input neurons have synaptic connections to all output neurons. *Source:* From W. Schneider. Connectionism. Is it paradigm shift for psychology? *Behavior Research Methods, Instruments & Computers,* Volume 19, pp. 73–83. Copyright © 1987. Reprinted by permission of Psychonomic Society, Inc.

learning. It has been argued that human cognition corresponds to patterns of firing over large numbers of neurons. The learning problem for such a network is to learn strengths of association among neurons such that when a particular pattern of activation occurs on the input neurons, a desired pattern of activation appears on the output neurons. Seen in terms of classical conditioning, each input neuron corresponds to a CS and each output neuron corresponds to a US. One proposal for how to achieve such association of patterns is connectionist learning rule, called the **delta rule**, which is based on the Rescorla–Wagner equation.

Figure 2.16 illustrates an application of this modeling approach to a medical diagnosis problem studied by Gluck and Bower (1988). Their subjects studied records of fictitious patients who suffered from four symptoms (a bloody nose, stomach cramps, puffy eyes, and discolored gums) and made discriminative diagnoses as to which of two hypothetical diseases the patients had. In this case, the four symptoms are the inputs and the two diseases are the outputs. For each patient, the input neurons corresponding to that patient's symptoms would be active and the goal would be to have the output neuron corresponding to that patient's disease active. This problem can be mapped onto the Rescorla–Wagner theory by having the inputs correspond to CSs and the outputs to USs. Just as a rat is trying to predict shock in a particular stimulus situation, subjects are trying to predict a disease given certain symptoms. The delta

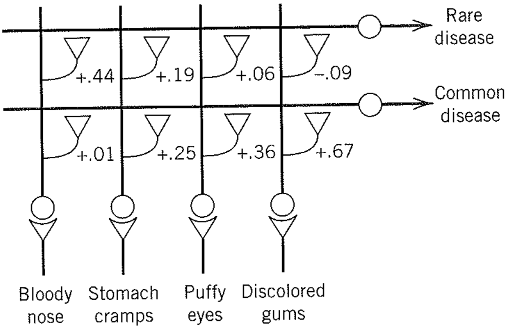

FIGURE 2.16 An adaptation of Figure 2.15 to the Gluck and Bower experiment in which subjects associate symptoms to diseases.

rule treats the neural learning as if each input-to-output association is being learned as a separate CS–US association. In the case of Figure 2.16, 4 × 2 = 8 associations are being learned.

In the context of neural modeling, the Rescorla–Wagner equation, or the delta rule, is used for adjusting the strength of the synaptic connection between the input neuron i and an output neuron j. The delta rule is stated as

$$\Delta A_{ij} = \alpha A_i (T_j - A_j)$$

where ΔA_{ij} is the change in the strength of synaptic connections between input i and output j; α is the learning rate; A_i is the level of activation of input neuron i; A_j is the level of activation of output neuron j; and T_j is the target or desired activity of j. This equation should be compared with the Rescorla–Wagner equation, which states

$$\Delta V = \alpha (\lambda - V)$$

where ΔV corresponds to ΔA_{ij}; α corresponds to αA_i; λ corresponds to T_j; and V corresponds to A_j. As in the Rescorla–Wagner theory, learning is proportional to the difference $T_j - A_j$. According to the delta rule, learning is also proportional to A_i, the level of activation of the input neuron. Recall that one proposal for an extension of the Rescorla–Wagner theory was to make learning proportional to stimulus salience.

In the case of a network, as shown in Figures 2.15 and 2.16, where there are m possible inputs and n possible outputs, the system uses the delta rule to simultaneously learn some $m \bullet n$ synaptic connections. The behavior of the whole system can be complicated, but it still reproduces the basic competitive learning behavior of the original Rescorla–Wagner theory.

The actual Gluck and Bower experiment was quite complex. Subjects saw hundreds of patient descriptions reflecting different combinations of symptoms. Each combination had a different probability of each disease. Overall, one disease was much rarer than the other, and each symptom had a differential association with the disease. The subjects were supposed to learn from this experience how to predict a disease given a pattern of symptoms. Figure 2.16 illustrates the final strengths of association that should be learned in accordance with the delta rule. These strengths of association did an excellent job of predicting subjects' classification behavior and how they rated each symptom as to whether it was diagnostic of the rare disease or the common disease.[5] As seen from the synaptic weights in Figure 2.16, a bloody nose was treated as more symptomatic of the rare disease (.44 strength versus .01 strength), whereas the other three symptoms were treated as more symptomatic of the common disease. In their ratings of the symptoms after the experiment, subjects agreed that only a bloody nose was predictive of the rare disease.

The delta rule has become an important construct in theories of neural learning. It has been used in a wide variety of models of neural processing and has been imported by computer science to build models of machine learning. As this example illustrates, it has also been used to predict complex human learning. At the same time, it has a number of difficulties similar to the difficulties that the Rescorla–Wagner theory has with classical conditioning. For instance, it cannot explain the learning of the configural cues. Gluck and Myers (1993) suggest that it is a good model of cortical learning but that other sorts of learning are controlled by a subcortical structure called the hippocampus. The next chapter discusses the role of the hippocampus in learning.

The Rescorla–Wagner theory corresponds to a popular theory of competitive learning among neural elements, called the delta rule.

Final Reflections on Classical Conditioning

Classical conditioning is a phenomenon defined by experimental procedures: the US is made contingent on the CS, and, as a consequence, the CS acquires the capacity to elicit a CR. Early in the history of classical conditioning, researchers tended to view the learning that was taking place as an unconscious and automatic consequence of the contiguity of CS and US. Classical conditioning was attractive in part because it was seen to embody pure and simple learning, uncontaminated by cognition on the part of the organism. Just as early researchers had hoped, classical conditioning studies on some animals, particularly simple animals such as the *Aplysia* that have no central nervous system, have provided insight into associative learning free of cognition. However, recent research has also demonstrated that classical conditioning in more complex organisms, including humans, frequently involves cognitive influences.

---

[5] Gluck and Bower's model actually has a single output, which varies from +1 to –1, depending on the probability of the rare disease. The model in Figure 2.15 is formally equivalent but more in keeping with the Rescorla–Wagner theory. In this model, each disease is predicted separately in a 0–1 scale, and the maximum strength of association is 1. The Gluck and Bower values can be determined by subtracting the strength of the common disease from the strength of the rare disease.

As discussed at the beginning of the chapter, different processes in the nervous system underlie different instances of classical conditioning. Despite these many differences, there are strong behavioral similarities across most forms of conditioning: similar acquisition and extinction histories, spontaneous recovery, generalization, temporal parameters, relationships between the CR and the UR, and so on. The reason for the similarities lies in the fundamental adaptiveness of classical conditioning. Classical conditioning allows organisms to respond adaptively in anticipation of a biologically significant UR. The common functionality of classical conditioning across species underlies its common behavioral properties. An analogy can be drawn from the relationship between the eye of the mammal and the eye of the octopus. The two have independent evolutionary histories and are formed from different tissue, but they function almost identically. They function so similarly because both have to deal with the same problem of extracting information from light. Likewise, classical conditioning is similar across organisms in terms of its behavioral properties because it serves the same function of allowing the organism to respond in anticipation of the US.

Thus, certain generalities at a behavioral level in classical conditioning may not be supported by generalities at a physiological level. In science diverse underlying processes often give rise to common higher-order generalities in the behavior of the system. For example, in economics industries producing very different products display the same economic realities. Similarly, in biology different organisms display similar relationships. For instance, as reviewed in Chapter 4, rather different creatures follow rather similar principles of foraging for food.

This book identifies many generalities in learning that apply to diverse species. The study of learning and memory has traditionally been about identifying these generalities and understanding them. Much of psychology is devoted to understanding generalities at the behavioral level. However, it should not be concluded that all species learn the same way at the behavioral level. Species show different preferences for various CS–US combinations, reflecting differences in their biological makeup and in their environment. The closest thing to a universal claim that can be made about classical conditioning is that it tends to reflect an adaptive response by the organism to CS–US contingencies. However, even this claim depends on carefully choosing the meaning of "adaptive." For instance, it is really not adaptive for chemotherapy patients to develop food aversions, although their tendency to form food aversions might be adaptive in other contexts, where it would lead them to avoid poisonous food.

The behavioral properties of classical conditioning reflect its adaptive character.

Further Readings

A number of texts are devoted to discussion of classical and instrumental conditioning, such as that by Domjan (1998). Wasserman & Miller (1997) provide a recent review of current research on classical conditioning. Miller, Barnet, and Grahame (1995) have written a review of the Rescorla–Wagner theory. The September 1992 issue of *Scientific American* was devoted to the relationship between brain and mind and presented a number of relevant articles, such as one by Kandel and Hawkins on the biological basis of learning and another by Hinton explaining recent developments in connectionist theories of learning. Among them is a mechanism called backprop, which is a popular extension of the delta rule. Thompson, Donegon, and Lavond (1988) published a fairly exhaustive review of the psychobiology of learning. Bernstein and Borson (1986) reviewed research on learned food aversions. Journals that frequently publish research on classical conditioning include the *Journal of Experimental Psychology: Animal Behavior Processes; Behavioral Neuroscience;* and *Animal Learning and Behavior.*